Reply to the Editor:

We are gratified by Dr Anderson's solid endorsement of our editorial, helping to clarify this rather confused and confusing situation surrounding the use of the word "actual" in the assessment of intrinsic valve performances. We fully agree with him regarding the importance of the concept of competing risks environment. His proposals on reporting long-term experiences with replacement heart valves are excellent, and we recommend that future authors follow his advice.

The letter by Grunkemeier, Takkenberg, and Jamieson incorrectly summarizes in its first paragraph the content of our editorial and therefore adds unnecessary confusion. Dr Bodnar was probably the first to introduce competing risks to cardiac surgery; Dr Blackstone and colleagues have used competing risks analyses extensively in multiple settings of adult and, particularly, congenital heart disease, and both will continue to do so to answer questions that methodology was designed to answer. The method is not the problem! The problem is its inappropriate use in answering questions related to intrinsic properties of heart valve substitutes, such as comparative durability. From its beginning, our editorial was clear about the specific, focused context of our remarks. Rather than restate the entire editorial to make 1 point clear, perhaps an analogy would be helpful.

Pretend we wish to compare durability (tread life) of 3 brands of tire. (We understand that, just like structural valve deterioration [SVD], this is a continuous process and not an event; for simplicity, however, we will estimate durability by the time to change tires because of excessive tread wear, just as we may estimate SVD by time to valve replacement.) Likely there are specific risk factors for tire durability, but let us imagine that miles traveled is the strongest (just as young age is the strongest correlate of SVD). Now, let us say that our study of tire durability for brand A is dominated by a fleet of cars operated by traveling salespersons, study of brand B by commuters living within 5 miles of work, and study of brand C by mothers with children involved in many after-school activities. After a stated period, most salespersons have had to change tires, a number of mothers have, but few short-distance commuters have. Should we conclude that A and C tires wear out too fast, and we should switch to brand B? If we instead compare the tires in a distance-specific fashion using time-to-event (actuarially based) analysis methods, we would be getting closer to a fair comparison of brands A, B, and C by isolating tire properties from the driving specifics of their owners. In the case of SVD, the only universally found risk factor is age of patient at implant, so age-specific durability provides a reasonable comparison of bioprosthetic device durability. As with the tires, estimates of durability are properly made by time-to-event-type analyses.

You may then ask: What tire should we recommend for the driving habits of a given person? The answer may well depend on how long different people intend to keep their cars. A person who leases a new car every 3 years no matter how little he or she drives may never need to change tires. Trading in a car, like death of a patient, is a competing risk. For someone who drives little, there is no reason to pay a premium for a tire of superior durability if at trade-in there is unlikely to be much tread wear! However, it is important to understand that the car-trading habits of owners do not themselves affect intrinsic properties of the tires. So, interesting as competing risks analysis is for answering some questions such as those posed, for simply comparing intrinsic durability of tires, these trading habits are extraneous and should not be allowed to confound the comparison.

Choice of a prosthesis for a given patient, like choice of a tire, may well depend on a number of factors, such as the patient's expected longevity, which itself has many correlations that are unrelated to intrinsic properties of the device. These are appropriately evaluated by competing risks analyses. However, longevity of the person, like car-trading habits, is a far more complex matter than the intrinsic properties of the prostheses. Estimates of whether or not a patient will live to experience SVD of their prosthesis will be different for every patient, despite the prosthesis retaining its unchanged intrinsic durability.

It is important to recognize that in *both* time-to-event and competing risks methods, time to SVD and time to death are assumed to be completely independent of one another. Thus, the fundamental "actuarial" calculation for each component of a competing risks analysis can be performed independently of one another, after which a mathematical combination is made. Thus, actuarial estimates are fundamental and unchanging, whereas competing risk estimates will vary, depending on what set of events is considered competing or ignored. (You may protest that it is irrational to think there is no linkage between SVD and death, and you are probably right. However, that is the assumption that both "actuarial" and competing risks methods assume. Methods for dealing with linked events, called methods for "informative censoring," are still not well developed.)

In conclusion, we suggest that Drs Grunkemeier, Takkenberg, Jamieson, and Miller need to realize the magnitude of the mistake they make. Actuarial analysis and cumulative incidence assessment are not *competing* but *complementary* methods. One is apple, the other orange, and oranges should not be blamed for not being apples. The question is not which is better—the question is which should be used for what purpose. Cumulative incidence cannot be used to assess let alone compare intrinsic valve properties. We were not advocating that articles employing competing risks analysis appropriately be "banned" from our journals but that research focused on long-term intrinsic performance of replacement heart valves and their comparison not be published with inappropriate analyses of competing risks.

Endre Bodnar, MD[a]
Eugene H. Blackstone, MD[b]
Editor in Chief[a]
The Journal of Heart Valve Disease
Northwood, UK
Cleveland Clinic[b]
Cleveland, Ohio

doi:10.1016/j.jtcvs.2006.03.070

Reporting "actual freedom" should not be banned

To the Editor:

The life of a bioprosthetic heart valve usually ends as a result of structural valve deterioration (SVD) or death of the patient. In this competing risks situation, the probability of SVD is estimated by the cumulative incidence function, sometimes referred to as "actual" analysis.

The editorial by Bodnar and Blackstone published in the January 2006 issue of the *Journal*[1] criticized this method and recommended its prohibition. Specifically,

Bodnar and Blackstone object to: (1) terminology: using the name "actual" instead of "cumulative incidence"; (2) presentation: plotting its complement and calling it "actual freedom"; and (3) comparison: using this statistic instead of the Kaplan–Meier (KM) method to compare valve performance. We disagree with Bodnar and Blackstone on all 3 points.

Regarding the first point, terminology, Bodnar and Blackstone state: "We should stop using the term 'actual freedom.' . . . There is no reason to abandon the expression 'cumulative incidence.'" Statisticians use several names for this method, including "cumulative incidence," "crude probability," "crude incidence," "cause-specific failure probability," "absolute cause-specific risk," and "subdistribution function." "Actual" is just a shortcut name, like saying "Ross"—not a medical term—instead of "pulmonary autograft" or "linearized rate"—not a statistical term—instead of "the constant hazard rate of an exponential distribution."

Bodnar and Blackstone also state: "Synonyms of 'actual' are current, eventual, and real, and the implication of its use is that it is more real than the actuarial estimate." That is exactly the point; the term "actual" is appropriate because it *is* "more real" than the KM actuarial estimate for the individual patient's perspective because it estimates the percentage of patients who will "actually" experience an event. If one has a strong emotional aversion to the use of the term "actual," then the statistical terms "subdistribution function" instead of "actual failure" and "subsurvival function" instead of "actual freedom"[2] can be substituted.

Regarding the second point, presentation, Bodnar and Blackstone state: "Other authors . . . displayed graphically the complement of cumulative incidence, which normally rises from zero to a certain positive value, so that the new curve declined from 100% to a certain value." Although the probability of having an event is the more direct concept, its complement—the probability of being event-free—contains the same information and is usually preferred in survival analysis and complication-free analysis.

In reference to the third point, comparison, Bodnar and Blackstone state: "Cumulative incidence . . . must not be used to define or compare valve performance. This should be done using actuarial [KM] methods." The first two issues are really a matter of taste, but this one is a matter of substance and statistical correctness. The subdistribution and subsurvival functions estimate the probability of having, or not having, respectively, SVD in the presence of the competing risk of death. As such, it would not be fair to use them to compare, say, valve A from a series with a high death rate to valve B from a series with a low death rate. If the death rates in 2 series were similar, however, then the comparison of SVD subdistribution or subsurvival functions would be appropriate.[3] But this same caveat applies to KM actuarial estimates. Comparing high SVD rates with valve X in a younger population to low SVD rates with valve Y in an older population would not be correct, as the difference could be due to patient age and other patient-related factors alone. To make a fair comparison of KM curves, the patient groups should also be similar. But there also is a deeper, technical issue involved here. The subdistribution function gives a proper probability of experiencing SVD; the KM estimate does not.

Bodnar and Blackstone also state: "We assume that patients who die before a nonfatal event occurs were just as likely, while they were alive, as anybody else to have experienced that event, even though they didn't. Now, that sounds more sensible, doesn't it?" That is true, if indeed the events of death and SVD are independent. It may sound sensible, but it cannot be demonstrated using competing risks data,[4] and for this reason the KM curve resulting from this assumption has dubious value. It has been widely criticized as "an incorrect use of the Kaplan–Meier method,"[3] "a meaningless quantity,"[5] "irrelevant,"[6] and "inappropriate for estimation purposes in the presence of competing risks, while the cumulative incidence is appropriate."[7] That is the statistical argument.[8] Then there is the clinical issue: The KM curve gives the probability of SVD in the counterfactual (or artificial) situation where death has been eliminated. This, of course, will never happen.

Bodnar and Blackstone conclude that "the Journal will no longer publish 'actual freedom' results in articles reporting long-term performance of replacement valves." Although this prohibition was stated to only apply to the reporting of long-term performance of heart valves, this method has other valuable uses in cardiothoracic clinical research. Examples include the analysis of nonfatal complications such as aortic reoperation or stroke after surgical repair of aortic dissection, reintervention after thoracic aortic stent grafting, recurrent obstruction after esophagectomy, complications after lung volume reduction operations, and transplant graft disease in cardiac allograft recipients. Presumably, these other clinical applications are not affected by the proposed prohibition. In light of the above, we argue that the use of "actual freedom" for reporting heart valve complications or those following other cardiothoracic interventions should not be banned from *The Journal of Thoracic and Cardiovascular Surgery*.

Gary L. Grunkemeier, PhD[a]
Johanna J. M. Takkenberg, MD, PhD[b]
W. R. Eric Jamieson, MD[c]
D. Craig Miller, MD[d]
Medical Data Research Center[a]
Providence Health System
Portland, Ore
Department of Cardio-Thoracic Surgery[b]
Erasmus University Medical Center
Rotterdam, The Netherlands
St Paul's Hospital–Heart Centre[c]
University of British Columbia
Vancouver, British Columbia, Canada
Department of Cardiothoracic Surgery[d]
Stanford University Medical School
Stanford, Calif

References

1. Bodnar E, Blackstone EH. An "actual" problem: another issue of apples and oranges. *J Thorac Cardiovasc Surg.* 2006;131:1-3.
2. Crowder M. Classical competing risks. Boca Raton: Chapman & Hall/CRC, 2001:200.
3. Lin DY. Non-parametric inference for cumulative incidence functions in competing risks studies. *Stat Med.* 1997;16:901-10.
4. Tsiatis A. A nonidentifiability aspect of the problem of competing risks. *Proc Natl Acad Sci.* 1975;72:20-2.
5. Klein JP, Andersen PK. Regression modeling of competing risks data based on pseudovalues of the cumulative incidence function. *Biometrics.* 2005;61:223-9.
6. Pepe MS, Mori M, Longton G, Pettinger M, Fisher LD, Storb R. Kaplan-Meier, marginal or conditional probability curves in summarizing competing risks failure time data? *Stat Med.* 1993;12:737-51.
7. Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: new representa-

tions of old estimators. *Stat Med.* 1999;18:695-706.
8. Grunkemeier GL, Jin R, Eijkemans MJC, Takkenberg JJM. Actual and actuarial probabilities of competing risks: apples and lemons. *Ann Thorac Surg.* 2007; in press.
doi:10.1016/j.jtcvs.2006.03.069

Atrial fibrillation surgery: Is it time to draw specific recommendations?

To the Editor:

We read with interest the article by Stulak and colleagues[1] evidencing the limits of radiofrequency (RF) ablation compared with those of the maze technique. The reproduction by RF of the lesion pattern featured in the maze operation aims to obtain comparable results with shorter ischemic times and reduced incision-related bleeding risks. In fact, the results reported by the authors cast a shadow on this assumption; freedom from atrial fibrillation (AF) appears less satisfactory in patients treated with ablation compared with that seen in patients undergoing the classical maze operation, both at hospital discharge and at a median 8-month follow-up.[1]

The RF-based technique addresses the same atrial lesion pattern as that of the maze operation; subsequently, the different outcome of the treated patients is likely to be related to the efficacy of RF in terms of realization of transmural lesions. The purpose of RF ablation (ie, "to create full Cox maze lesions")[1] is challenged.

The creation of transmural lesions is the aim of ablative surgery and should be considered a "must," ensuring procedural effectiveness. Unfortunately, intraoperative confirmations that this goal has been achieved are not available with the current methodologies, leaving a question mark on this issue, which is recalled by the results of Stulak and colleagues.[1]

Nevertheless, the widespread adoption of other-than-maze techniques (mostly based on the topical application of various energy sources to provoke conduction blocks) is a measure of how uncomfortable cardiac surgeons generally are with the maze operation. It also explains why the maze operation is adopted in few institutions of excellence and performed in selected patients instead of being a standardized and accepted procedure.

In fact, in daily practice cardiac surgeons are favoring less-invasive, more easily reproducible, and "softer" approaches, whose technology is in continuous evolution and whose efficacy is currently under evaluation.

The absence of guidelines or consensus statements, specifically those focusing on AF surgery, has been underlined by others.[2,3] We believe that the extensive information gained from the clinical application of both the maze operation (since the early 1990s) and energy-based ablative surgery (since the late 1990s) would allow the cardiovascular community to draw recommendations indicating to surgeons when and how to proceed with AF surgery in patients undergoing cardiac operations. We are convinced that such recommendations would be fundamental to orientate both the cardiac surgeons in their daily practice and the related research in its evolution.

Francesca di Marco, PhD, MD
Gino Gerosa, MD
Department of Cardiological, Thoracic, and Vascular Sciences
Padua University Medical School
Division of Cardiac Surgery
Padova, Italy

References

1. Stulak JM, Dearani JA, Sundt TM 3rd, Daly RC, McGregor CG, Zehr KJ, et al. Superiority of cut-and-sew technique for the Cox maze procedure: comparison with radiofrequency ablation. *J Thorac Cardiovasc Surg*. 2007;133:1022-7.
2. Shemin RJ, Cox JL, Gillinov AM, Blackstone EH, Bridges CR, Workforce on Evidence-Based Surgery of the Society of Thoracic Surgeons. Guidelines for reporting data and outcomes for the surgical treatment of atrial fibrillation. *Ann Thorac Surg*. 2007;83:1225-30.
3. Khargi K. Why do we need an arrhythmia innovation center? Available at: http://www.ctsnet.org/sections/innovation/arrhythmia/articles/article-1.html. Accessed May 2007.
doi:10.1016/j.jtcvs.2007.05.070

Reply to the Editor:

We agree with Drs di Marco and Gerosa that it is time to synthesize the large amount of data on surgical treatment of atrial fibrillation (AF), conflicting as it might be, and formulate guidelines for clinicians. Fortunately, 2 recent publications address this need. The Society of Thoracic Surgeons Workforce on Evidence-based Surgery has recently published "Guidelines for reporting data and outcomes for the surgical treatment of atrial fibrillation."[1] This group proposes standard descriptions of preoperative AF, of the surgical procedure, and of the lesion set performed. Furthermore, the guidelines propose uniform reporting of postoperative protocols, follow-up methodologies, and outcome rhythm.

A second publication by the Heart Rhythm Society Task Force on Catheter and Surgical Ablation of Atrial Fibrillation suggests the following indications for surgical treatment of AF: (1) symptomatic patients with AF undergoing other cardiac surgical procedures; (2) selected asymptomatic patients with AF undergoing cardiac surgery in whom the ablation can be performed with minimal risk; and (3) symptomatic patients with AF who prefer a surgical approach, have experienced 1 or more failed attempts at catheter ablation, or are not candidates for catheter ablation.[2]

Although surgical intervention for AF has been performed for 2 decades, prospective multicenter clinical trials are still needed to define the relative safely and efficacy of various surgical tools and techniques. In addition, surgeons should use consistent definitions of procedural success and follow-up methodology to compare the success of various surgical methods, as well as outcomes of surgical intervention, catheter ablation, and medical treatment.

John Stulak, MD
Hartzell V. Schaff, MD
Department of Cardiovascular Surgery
Mayo Clinic College of Medicine
Rochester, Minn

References

1. Shemin RJ, Cox JL, Gillinov AM, Blackstone EH, Bridges CR. Workforce on Evidence-Based Surgery of the Society of Thoracic Surgeons. Guidelines for reporting data and outcomes for the surgical treatment of atrial fibrillation. *Ann Thorac Surg*. 2007;83:1225-30.
2. European Heart Rhythm Association (EHRA), European Cardiac Arrhythmia Society (ECAS), American College of Cardiology (ACC), American Heart Association, Society of Thoracic Surgeons (STS), Calkins H, et al. HRS/EHRA/ECAS expert Consensus Statement on catheter and surgical ablation of atrial fibrillation: recommendations for personnel, policy, procedures and follow-up. A report of the Heart Rhythm Society (HRS) Task Force on catheter and surgical ablation of atrial fibrillation. *Heart Rhythm*. 2007;4:816-61.
doi:10.1016/j.jtcvs.2007.07.058